

Image Processing Method,
Device and Storage Medium Therefor

BACKGROUND OF THE INVENTION

5 Field of the Invention

The present invention relates to discriminating orientation of an input image.

Related Background Art

10 In recent years, document digitizing devices like character recognition devices or facsimile apparatuses have been developed to be in a realization as computer technology has higher performance. With the wide spread of these devices, principally in business applications, demand has been growing for converting
15 vast amounts of documents to image data at a time with a scanner having an auto document feeder (ADF). For that purpose, efficiency is attracting the most attention from industry.

Typically, ADFs are fed with numerous documents of
20 the same size and orientation. If documents in longitudinal direction and lateral direction are mixed, obtained image data will have different orientations. If some documents have different sizes, such documents may tend to rotate in travelling in an ADF and obtained
25 image data will be more likely to have a wrong orientation.

Since a conventional character recognition device

or a filing device assumes that documents is processed
in a common orientation, an operator arranges their
orientations to be in a common orientation before the
documents are fed into such a device. Moreover, in
5 order to ensure proper operations, after documents have
been processed by a device, an operator must visually
inspect all document images on a display one after
another and if they have wrong orientations, the
operator must provide an instruction to arrange it by
10 an appropriate angle (90°, 180° or 270°) accordingly.
Therefore, processing vast amounts of documents at a
time may significantly degrade practicability of the
devices in the point of light efficiency and burden to
operators.

15

SUMMARY OF THE INVENTION

The purpose of the present invention is to reduce
operators' burdens and improve efficiency by
automatically discriminating orientations of a
20 plurality of images which are not always fed in a
common orientation.

Another purpose of the present invention is to
reduce possible burdens to operator and improve
efficiency by automatically discriminating orientation
25 of each of a plurality of images which may have
different orientations.

Furthermore, another purpose of the present

invention is to improve accuracy of processing by judging if results of automatic discrimination are not normal and informing exteriors of the results.

5 BRIEF DESCRIPTION OF THE DRAWINGS

FIG. 1 is a functional block diagram according to the present invention.

FIG. 2 is a flow chart for showing cooperation between a processes of a document image orientation
10 correction part 4 and a correction instruction part 6.

FIG. 3 shows an example of an input window for document image orientation correction instructions.

FIG. 4 shows a hardware configuration of devices according to the present invention.

15 FIG. 5 is a flow chart for showing a process of a correction instruction part.

FIG. 6 is a flow chart for showing a process when an orientation correction is in Auto mode.

20 DETAILED DESCRIPTION OF THE PREFERRED EMBODIMENTS

FIG. 4 shows a hardware configuration of a device according to the present invention.

In FIG. 4, the numeral 401 shows a central processing unit (CPU), which controls processes
25 according to the present invention by executing control program stored in memory 402. Steps shown in a flow chart described below are also executed by CPU 401.

Memory 402 is a storage device consisting of RAMs, ROMs, and hard disks, in which control programs for the CPU 401 and various parameters, input image data, and character recognition dictionaries are stored. The numeral 403 shows an external storage medium which can be attached to and removed from the present device, such as optical disks, magnetic disks, magneto-optical disks, and magnetic tapes. Programs and data to be stored in memory 402 are read out from this external storage medium. Any result of processes is provided to external storage medium 403. The numeral 404 shows a communication I/F for communicating data to other terminals through a network or public circuit and programs and data to be stored in memory 402 and the process result may be provided through this communication I/F. The numeral 405 shows an input means such as a keyboard or pointing device, which is used to transmit the operator's instructions. Instructions to read images, to start character recognition, and to specify destination where text data resulting from a recognition process is stored are entered through input means 405. The numeral 406 shows a scanner which optically reads a document and provides a document image as electrical signals to the present device. The numeral 407 shows a display such as a CRT or liquid crystal display, which is used to display text data resulting from process as well as to be an

interface for an operator. The numeral 408 shows a laser beam printer (LBP) or an ink jet printer, which is used to print out text data resulting from process in desired fonts.

5 FIG. 1 is a functional block diagram according to the present invention.

 In FIG. 1, reference numeral 2 shows an input part for entering document images, which are fed into the present device from scanner 406 or external storage
10 medium 403 or from other terminals through communication I/F 404. The numeral 4 shows a document image orientation correction part for correcting the orientation of an input document image, 6 is a correction instruction part, 8 is an area
15 discrimination part for document images, and 10 is a character recognition part for recognizing characters, all of which may be implemented by CPU 401 as shown in the flow chart described below. The numeral 12 shows an output part for providing recognition results to
20 display 407, printer 408, or external storage medium 403, or to other terminals through communication I/F 404. The numeral 14 shows a document image memory which may be implemented by memory 402.

 An operation of devices according to the present
25 invention is now described below.

 A document is first converted to a predetermined image through input part 2 and stored in document image

memory 14, regardless of mediums which contains the documents. Acquired document images are provided to document image orientation correction part 4 through document image memory 14 and whether any correction is required or not depends on an instruction from correction instruction part 6. If correction instruction part 6 does not indicate an instruction to correct the image, document image orientation correction part 4 does not perform any correction and then entered image of orientation is provided to area discrimination part 8. An operation of document image orientation correction part 4 and correction instruction part 6 will be described later in detail. Results of image correction performed by document image orientation correction part 4 is again accumulated in document image memory 14. Area discrimination part 8 analyzes images to divide it into subareas, depending its attribute such as character, table, drawing, image, and photograph, and then provides to memory 402 block data for each subarea which consists of the number assigned to area in order of area acquisition or closeness to a reference position in the coordinate system, the attribute, the size, and the position information. The block data is stored in memory 402. Image data within in a subarea having the attribute "character" which is determined to be a character area by area discrimination part 8 is transmitted to

character recognition part 10 for character
recognition. Text data resulting from the character
recognition is transmitted to output part 12 and then
provided as an output through display 407 or printer
5 408 in fonts corresponding to graphic character codes
for the text data. During outputting the text, layout
of an original document may be reproduced by arranging
appropriate fonts within a character area where the
text data was recognized based on the position
10 information for the area contained in the block data.
Alternatively, the text data resulting from the
character recognition or both text and block data may
be stored in external storage medium 403 or provided to
other terminals through the communication I/F.

15 Now, document image orientation correction part 4
and correction instruction part 6 will be described
below in detail.

FIG. 2 is a flow chart for showing cooperation
between processes of document image orientation
20 correction part 4 and correction instruction part 6.

Step S202 is a routine to input a document image
orientation correction instruction, wherein an operator
uses input means 405 to specify a desired mode. The
mode specified in this step, that is, the instruction
25 as to whether the document image orientation should be
automatically corrected is stored in memory 402. FIG.
3 shows an example of instruction input window actually

displayed on display 407. When the checkbox for the instruction "Auto" (300) in the group "Rotation" on this window is marked, automatic orientation correction will be performed by document image orientation correction part 4. When the checkbox for the instruction "None" is marked, the automatic orientation correction will not be performed. It should be appreciated that marking of such an instruction checkbox is a signaling image to allow the operator to visually know whether instruction is selected and that the selection can be effected by placing a mouse cursor on instruction checkboxes and clicking a mouse button on it. For each of the groups "Rotation", "Black-and-White Inversion", "Tilt Correction", "Insert Page", and "Input Image", only one checkbox can be marked and if any one of the checkboxes within a group has been marked as described above, the other checkboxes within the group can not be marked, that is, the instructions associated with those checkboxes can not be selected. The selected instruction within each group is stored in memory 402 as flag data associated with the group data. Such flag data will be updated every time the mouse button is clicked. Even if the instruction "Auto" has been selected for the group "Rotation", the instruction "Auto" should be deselected when any of the other instructions "None", "90° CW", "180°" and "90° CCW" is selected. All the instructions

selected on the instruction input window shown in FIG.
3 are stored in memory 402 as setting data for image
input.

Step S204 is a routine to actually input a
5 document, wherein, an instruction to read the document
is issued to the driver for a scanner, if used, or
document data read from a file is converted to a
predetermined image and in either case, document data
is stored in document image memory 14 in an expanded
10 form.

Step S206 is a routine to decide subsequent
process, depending on the instructions selected in step
S202, wherein if it is determined from the setting data
for image input stored in memory 402 that the document
15 image orientation should be automatically corrected,
the process proceeds with step S208. If it is
determined in step S206 that the instruction "None" has
been specified for rotation, the control is passed to
step S212. If the instruction "90° CW", "180°" or "90°
20 CCW" has been specified, document image processed in
step S204 is rotated as specified and then control is
passed to step S212. Rotation process is identical to
that for step S210.

Step S208 is a routine to recognize orientation of
25 document images based on the data stored in memory in
step S204. In this step, character areas in document
image are extracted for each language and then

character patterns contained in each character area are rotated in the directions of 0°, 90°, 180°, and 270° so that a predetermined number of character patterns can be recognized by using appropriate dictionaries or in
5 an appropriate manner to that language. From the recognition results, that of highest average similarity among the four directions is determined to be a reference orientation. For this purpose, an algorithm for determining the angle for portrait format
10 orientation (0°, 90°, 180° or 270°) is assumed (see Japanese Patent Application No. 10-147620).

Step S209 is a routine wherein from the number of characters detected in the document image during character pattern extraction performed in step S208 as
15 a pre-process for recognizing document image orientation, it is determined whether orientation recognized in step S208 is correct. Such a determination is accomplished by comparing the number of characters in the document image with a
20 predetermined value previously stored in memory 402. If the number of characters in document image exceeds the predetermined value and therefore, it is determined that the orientation is correct, the process proceeds with step S210. Otherwise, that is, if the number of
25 characters in the document image is less than the predetermined value, the control is passed to the step S211. This routine detects beforehand any document

image containing much fewer characters, which may make it difficult to recognize the document image orientation.

5 Step S210 is a routine to rotate the document image stored in document image memory 14 to the reference orientation and to store corrected document image in document image memory 14 again on the basis of a rotation angle obtained at step S208.

10 Step S211 is a routine to give an alarm to an operator if it is determined in step S209 that the orientation discrimination for document image is not normal. As such an alarm, messages or signals is provided through display 407 or printer 408 or through a speaker (not shown); indicating that the document
15 image orientation can not be automatically corrected.

 Alternatively, in step S211, data which indicates that document image orientation could not be automatically corrected may be stored in memory 402 in correspondence to document images or data for the
20 document image which could not be automatically corrected can be provided from the output part for further determination in response to the instructions specified by the operator through keyboard 405 during output process of character recognition results.

25 Step S212 is a routine wherein area discrimination part 8 discriminates area of the document image stored in document image memory 14. This also includes a

pre-processing routine for area discrimination. Since area discrimination of this step is performed on image stored in document image memory 14, that image would have been rotated if an original document image has
5 been rotated in step S210.

Step S214 is a routine to recognize character patterns contained in any character area of document image data stored in document image memory 14 according to result of area discrimination performed in step S212
10 through character recognition part 10.

Step S216 is a routine to convert text data resulting from the character discrimination of the step S214 to a predetermined form so that the data can be supplied as an output through output part 12.

15 As described above, according to the present invention, any document image can be automatically corrected in correct orientation during document reading by incorporating the novel document orientation recognition technique into a character recognition
20 device through a user-friendly interface. Therefore, efforts required for confirmation before and after reading documents can be greatly reduced and a series of processes from document reading to the recognized character output can be performed at high precision.

25 In routine S209 described above, reliability of any orientation recognition result is determined based on the number of characters contained in the document.

It should be appreciated, however, that this determination is not limited to the discrimination described for this embodiment and that if the maximum similarity resulting from the character recognition performed during the orientation recognition is less than a predetermined value (a reference value previously stored in memory 402), it may be determined that a result of orientation recognition has a lower reliability and a process may proceed with step S211 to give an alarm indicating that the correction could not be performed.

In routine S211 described above, an alarm is given if it is determined that any result of orientation recognition has a lower reliability. It should be appreciated, however, that the alarm is not limited to that described for this embodiment and that for example, after process ends in step S216, an image file containing the original document image stored in document image memory 14, text data resulting from the character recognition, and block data may be stored together with an identification index to indicate that orientation could not be recognized.

Now, process for the case where the "Auto" mode is specified for the "Rotation" function will be described below.

FIG. 5 shows the process performed by correction instruction part 6. The term "orientation correction"

used hereinafter means automatic correction to be performed when an input document image is tilted at the angle of 90°, 180° or 270°, while the term "tilt correction" means another automatic correction to be performed when an input document image is tilted at a certain angle within $\pm 45^\circ$.

Step S502 is a routine to access an input window for document image orientation correction instructions and to prompt an operator to input desired instructions. The operator will specify desired modes of operation by marking appropriate checkboxes on the instruction input window as shown in FIG. 3.

Step S504 is a routine to determine whether an orientation correction instruction has been specified in the instruction input window invoked in step S502, that is, the "Auto" mode has been selected for the "Rotation" function as shown in FIG. 3. If the "Auto" mode has been selected, process proceeds with step S506. Otherwise, control is passed to step S514 labelled "Other Processes".

Step S506 is a routine to turn on orientation correction flags in memory 402 when it is determined that orientations should be corrected.

Step S508 is a routine to mark the checkboxes for the instruction "Auto" for the "Rotation" function in the window as shown in FIG. 3 in response to turn-on operation of step S506.

Step S510 is a routine to turn on also tilt correction flags when it is determined in step S504 that orientations should be corrected. In other words, if the "Auto" mode has been selected for the "Rotation" function, the "Tilt Correction" function will be automatically turned on without a specific instruction from the operator.

Step S512 is a routine to mark the checkboxes for the instruction "ON" for the "Tilt Correction" function in the window as shown in FIG. 3 in response to the turn-on operation of step S510.

Step S514 is a routine to perform other processes required to finish instruction inputs.

FIG. 6 is a flow chart for showing cooperation between processes performed by the correction control part, the document image orientation correction part, and the document image tilt correction part, respectively.

Step S602 is a routine to read out orientation correction flags and tilt correction flags from the memory, both of which have been specified through the correction instruction input part.

Step S604 is a routine to actually input a document, wherein an instruction to read documents is issued to a driver for a scanner, or document data read from a file is converted to a predetermined image and in either case, the document data of the input document

is stored in memory 402 in an expanded form.

Step S606 is a routine to determine whether the orientation correction flag read out in step S602 is ON or OFF. If the flag is ON, control is passed to step
5 S608 labelled "Document image orientation correction", and if the flag is OFF, process proceeds with step S610.

Step S608 is a routine to actually correct document image orientations. In this step, character
10 areas in the document images are extracted for each language and then character patterns contained in each character area are rotated in directions of 0°, 90°, 180° and 270° for character recognition. From the recognition results, highest average similarity among
15 the four directions is determined to be a reference orientation and then direction of reference orientation (0°, 90°, 180° or 270°) is provided to correct the document image orientations accordingly.

Step S612 is a routine to determine whether the
20 document orientation has been corrected properly. The control is passed to step S622 labelled "Alarm unable orientation correction" if the correction was not proper, or to step S614 labelled "Document image tilt correction" if the correction was proper. In this
25 step, accuracy for angle detection is determined based on the number of characters detected in the documents. That is, it will be determined in step S612 that

correction was proper if the number of document images having a similarity more than a certain threshold is equal to or greater than a predetermined value after the orientation has been corrected in step S608. It should be appreciated, however, that the determination as to whether the correction was proper is not limited to that described for this embodiment and that it may be determined proper if maximum average similarity among those obtained for the four angles in step S608 exceeds a predetermined value. Such predetermined values used for determination have been previously stored in memory 402.

Step S610 is a routine to determine whether tilt correction flags read out in step S602 is ON or OFF. The control is passed to step S614 labelled "Document image tilt correction" if the flag is ON, or to the step S626 labelled "Area discrimination" if the flag is OFF.

Step S614 is a routine to actually correct the document image tilt. This correction may be accomplished by using any well-known method.

Step S616 is a routine to determine whether the document tilt has been corrected properly. The control is passed to step S618 labelled "Alarm unable tilt correction" if the correction was not proper, or to step S626 labelled "Area discrimination" if correction was proper.

Step S618 is a routine to activate an alarm window on display 407 so that the operator can know that the tilt correction was not proper.

5 Step S620 is a routine to forcefully terminate the process, for example, by closing the alarm window.

Step S622 is a routine to activate an alarm window on display 407 so that an operator can know that orientation corrections were not proper.

10 Step S624 is a routine to forcefully terminate process, for example, by closing the alarm window.

Step S626 is a routine to discriminate any document image area of the data stored in the document image memory. This also includes a preprocessing routine for area discrimination.

15 Step S628 is a routine to recognize characters based on the area division result of step S626 and document image data stored in the memory.

20 Step S630 is a routine to convert character recognition results of step S628 to a predetermined form.

It should be appreciated that if it is determined in step S612 or S616 that corrections were not proper, an identification index to indicate the determination may be stored together with the image file. In this case, an alarm required in step S622 or S618 may be omitted.

25